Sequences with Minimal Time-Frequency Uncertainty

Reza Parhizkar*, Yann Barbotin*, Martin Vetterli

School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Abstract

A central problem in signal processing and communications is to design signals that are compact both in time and frequency. The variance is accepted as a good measure of compactness, and with this definition Heisenberg states that a given function cannot be arbitrarily compact both in time and frequency, defining an "uncertainty" lower bound. In continuous-time, Gaussian functions reach this bound. For sequences, however, this is not true; it is known that Heisenberg's bound is generally unachievable. For a chosen frequency variance, we formulate the search for maximally compact sequences as an exactly and efficiently solved convex optimization problem, thus providing a sharp uncertainty principle for sequences. Interestingly, the optimization formulation also reveals that maximally compact sequences are derived from Mathieu's harmonic cosine function of order zero. We further provide rational asymptotic expansions of this sharp uncertainty bound.

Keywords: Compact sequences, Heisenberg uncertainty principle, Time Spread, Filter design, Mathieu function, Discrete sequences, Circular statistics, Semi-definite relaxation

1. Introduction

Suppose you are asked to design filters that are sharp in the frequency domain and at the same time compact in the time domain. The same problem is posed in designing sharp probing basis functions with compact frequency characteristics. In order to formulate these problems mathematically, we need to have a correct and universal definition of compactness and clarify what we mean by saying a signal is spread in time or frequency.

These notions are well defined and established for continuous-time signals [1, 2] and their properties are studied thoroughly in the literature. For a continuous-time signal, we can define the time and frequency characteristics of a signal as in Table 1. Note the connection of these definitions with the mean and variance of a probability distribution function $|x(t)|^2/\left\|x\right\|^2$. The value of Δ_t^2 is considered as the spread of the signal in the time domain while $\Delta_{\omega_c}^2$ represents its spread in the frequency domain. We say that a signal is compact in time (or frequency) if it has a small time (or frequency) spread.

The Heisenberg uncertainty principle [1] states that continuous-time signals cannot be arbitrarily compact in both domains. Specifically, for any $x(t) \in L^2(\mathbb{R})$,

$$\eta_c = \Delta_t^2\, \Delta_{\omega_c}^2 \geq \frac{1}{4}\,, \tag{1}$$

where the lower bound is achieved for Gaussian signals of the form $x(t) = \gamma e^{-\alpha t^2}$, $\alpha > 0$ [2]. The subscript c stands for continuous-time definitions. We call η_c the *time-frequency spread* of x.

☆This work was supported by ERC Advanced Grant–Support for Frontier Research–SPARSAM, Nr: 247006.
*Corresponding author, Phone numbers: +41-21-693-5636 (Reza Parhizkar), +41-21-693-8146 (Yann Barbotin)
Email addresses: reza.parhizkar@epfl.ch (Reza Parhizkar), yann.barbotin@epfl.ch (Yann Barbotin), martin.vetterli@epfl.ch (Martin Vetterli)

domain	center	spread
time	$\mu_t = \frac{1}{\|x\|^2}\int_{t\in\mathbb{R}} t\|x(t)\|^2 dt$	$\Delta_t^2 = \frac{1}{\|x\|^2}\int_{t\in\mathbb{R}} (t-\mu_t)^2\|x(t)\|^2 dt$
frequency	$\mu_{\omega_c} = \frac{1}{2\pi\|x\|^2}\int_{\omega\in\mathbb{R}} \omega\|X(\omega)\|^2 d\omega$	$\Delta_{\omega_c}^2 = \frac{1}{2\pi\|x\|^2}\int_{\omega\in\mathbb{R}} (\omega-\mu_\omega)^2\|X(\omega)\|^2 d\omega$

Table 1: Time and frequency centers and spreads for a continuous time signal $x(t)$.

domain	center	spread
time	$\mu_n = \frac{1}{\|x\|^2}\sum_{n\in\mathbb{Z}} n\|x_n\|^2$	$\Delta_n^2 = \frac{1}{\|x\|^2}\sum_{n\in\mathbb{Z}} (n-\mu_n)^2\|x_n\|^2$
frequency	$\mu_{\omega_\ell} = \frac{1}{2\pi\|x\|^2}\int_{-\pi}^{\pi} \omega\|X(e^{j\omega})\|^2 d\omega$	$\Delta_{\omega_\ell}^2 = \frac{1}{2\pi\|x\|^2}\int_{-\pi}^{\pi} (\omega-\mu_\omega)^2\|X(e^{j\omega})\|^2 d\omega$

Table 2: Time and frequency centers and spreads for a discrete time signal x_n as extensions of Table 1 [2].

Although the continuous Heisenberg uncertainty principle is widely used in theory, in practice we often work with discrete-time signals (e.g. filters and wavelets). Thus, equivalent definitions for discrete-time sequences are needed in signal processing. In the next section we study two common definitions of center and spread available in the literature.

1.1. Uncertainty principles for sequences

An obvious and intuitive extension of the definitions in Table 1 for discrete-time signals is presented in Table 2, where

$$X(e^{j\omega}) = \sum_{n\in\mathbb{Z}} x_n e^{-j\omega n} \quad \omega \in \mathbb{R}\,, \tag{2}$$

is the discrete-time Fourier transform (DTFT) of x_n.

Using the definitions in Table 2 [2], we can also state the Heisenberg uncertainty principle for discrete-time signals as

$$\eta_\ell = \Delta_n^2\,\Delta_{\omega_\ell}^2 > \frac{1}{4}\,, \quad x_n \in \ell^2(\mathbb{Z}) \text{ with } X(e^{j\pi}) = 0\,, \tag{3}$$

where the subscript ℓ stands for *linear* in reference to the definition of the frequency spread. Note the extra assumption on the Fourier transform of the signal in (3). This assumption is necessary for the result to hold.

Example 1. Take $x_n = \delta_n + 7\delta_{n-1} + 2\delta_{n-2}$. It is easy to verify that that $\left|X(e^{j\pi})\right| = 0.22 \neq 0$, which violates the condition $X(e^{j\pi}) = 0$. The linear time-frequency spread of this signal according to Table 2 is $\eta_\ell = 0.159 < 1/4$.

In addition to the restriction on the Heisenberg uncertainty principle, the definitions in Table 2 do not capture the periodic nature of $X(e^{j\omega})$ for the frequency center and spread. In the search for more natural properties, we can adopt definitions for circular moments widely used in quantum mechanics [3] and directional statistics [4].

For a sequence x_n, $n \in \mathbb{Z}$, with a 2π periodic DTFT, $X(e^{j\omega})$ as in (2), the *first trigonometric moment* is defined as [5, 6]

$$\tau(x) = \frac{1}{2\pi\,\|x\|^2}\int_{-\pi}^{\pi} e^{j\omega}|X(e^{j\omega})|^2 d\omega\,. \tag{4}$$

The first trigonometric moment was originally defined for probability distributions on a circle. With proper normalization, this definition applies also to periodic functions.

domain	**center**	**spread**
time	$\mu_n = \frac{1}{\|x\|^2}\sum_{n\in\mathbb{Z}} n\|x_n\|^2$	$\Delta_n^2 = \frac{1}{\|x\|^2}\sum_{n\in\mathbb{Z}}(n-\mu_n)^2\|x_n\|^2$
frequency	$\mu_{\omega_p} = 1-\tau(x)$	$\Delta_{\omega_p}^2 = \frac{1-\|\tau(x)\|^2}{\|\tau(x)\|^2} = \left\|\frac{\|x\|^2}{\sum_{n\in\mathbb{Z}} x_n\, x_{n+1}^*}\right\|^2 - 1$

Table 3: Time and frequency centers and spreads for a discrete time signal x_n using circular moments, where $\tau(x)$ is defined in (4).

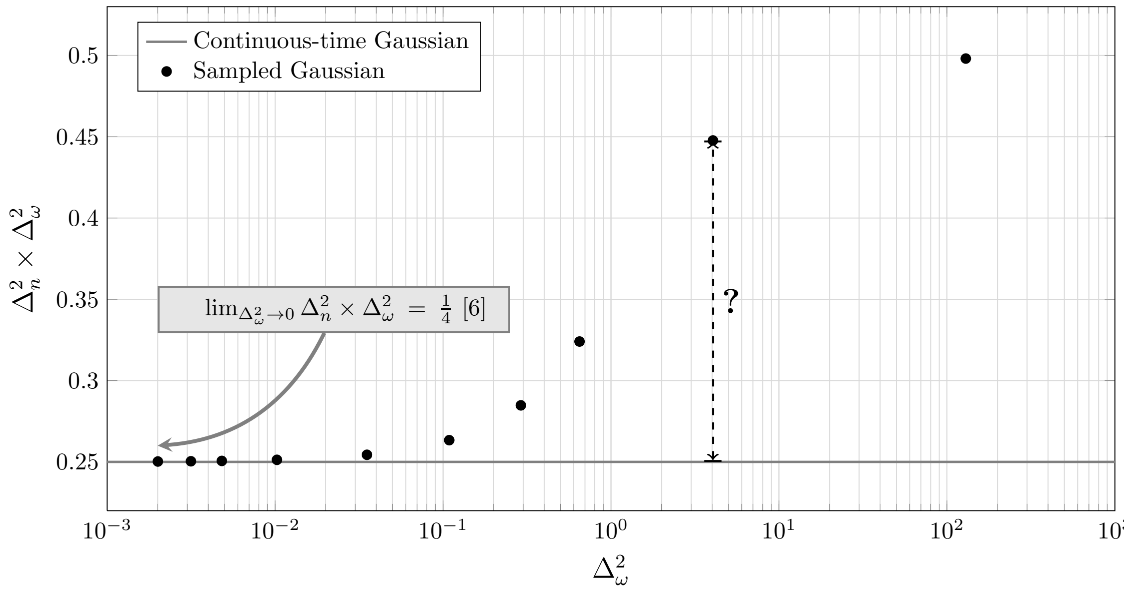

Figure 1: **Time-frequency spread of continuous vs. sampled Gaussians.** The solid line shows the 1/4 Heisenberg bound which is achieved by continuous Gaussian signals (with definitions in Table 1), while the markers show the time-frequency spread (according to Table 3) of sampled Gaussian sequences. The question is if the gap between the two curves is inherited from the properties of sequences or sampled Gaussians are not optimal in discrete domain.

Using (4), the *periodic frequency spread* is defined as [3]:

$$\Delta_{\omega_p}^2 = \frac{1-|\tau(x)|^2}{|\tau(x)|^2} = \left|\frac{\|x\|^2}{\sum_{n\in\mathbb{Z}} x_n\, x_{n+1}^*}\right|^2 - 1\,. \tag{5}$$

The definition of Δ_n^2 remains unchanged as in Table 2. These definitions are summarized in Table 3.

1.2. Contribution

In this paper, using the definitions in Table 3, we revisit the Heisenberg uncertainty principle for discrete-time signals.

We address the fundamental yet unanswered question: If someone asks us to design a discrete filter with a certain frequency spread ($\Delta_{\omega_p}^2$ fixed), can we return the sequence with minimal time spread Δ_n^2? In other words, the problem is to find the solution to

$$\begin{aligned} \Delta_{n,\mathrm{opt}}^2 = \underset{x_n}{\text{minimize}} \quad & \Delta_n^2 \\ \text{subject to} \quad & \Delta_{\omega_p}^2 = \text{fixed}\,. \end{aligned} \tag{6}$$

In order to provide an insight on the uncertainty principle in the discrete-time domain, we do a simple test. In figure 1 we show the time-frequency spread of continuous Gaussian signals by the solid line on $1/4$. This is the Heisenberg's uncertainty bound which is achievable by continuous-time Gaussians. Further, using the definitions in Table 3, we compute the time-frequency spread of sampled Gaussian sequences. As stated before, the time-frequency spread tends to $1/4$ as the frequency spread of Gaussians decreases, but when the frequency spread is large, the values are far from the uncertainty bound. Two questions arise here; is the $1/4$ uncertainty bound also tight for discrete sequences? and are sampled Gaussians the minimizers of the uncertainty in the discrete domain. Answers to these questions will be apparent if we can solve (6). We call the solution of (6) a *maximally compact* sequence.

Framing the design of maximally compact sequences as an optimization problem, we show that contrary to the continuous case, it is not possible to reach a constant time-frequency lower bound for arbitrary time or frequency spreads. We further develop a simple optimization framework to find maximally compact sequences in the time domain for a given frequency spread. In other words, we provide in a constructive and numerical way, a *sharp uncertainty principle for sequences*, later shown in Figure 4. We also show that the Fourier spectra of maximally compact sequences are in fact a very special class of Mathieu functions. Using the asymptotical expansion of these functions, we develop closed-form bounds on the time-frequency spread of maximally compact sequences.

1.3. Related Work

The classic uncertainty principle [1] assumes continuous-time non-periodic signals. Several works in the signal processing community also address the discrete-time/discrete-frequency case [7, 8, 9]. Our work bridges these two cases by considering the discrete-time/continuous-frequency regime.

Note that not all studies about the uncertainty principle concern the notion of spread. For example, the authors in [8] propose the uncertainty bound on the information content of signals (entropy) and [9] provides a bound on the non-zero coefficients of discrete-time sequences and their discrete Fourier transforms.

The discrete-time/continuous-frequency scenario has been recently encountered in many practical applications in signal processing. Examples include uncertainty principle on graphs [10] and on spheres [11]. Sudies on the periodic frequency spread can be found in [3] and [12]. The most comprehensive work on the uncertainty relations for discrete sequences is found in [6]. The authors show that $1/4$ is a lower-bound on the time-frequency spread, which can only be achieved asymptotically as the sequence spreads in time. We provide sharp achievable bounds in the non-extreme case which match the results in [6] in asymptotic.

This problem is similar to—although different than—the design of Slepian's Discrete Prolate Spheroidal Sequences (DPSS's). First introduced by Slepian in 1978 [13], DPSS's are sequences designed to be both limited in the time and band-limited in the frequency domains. For a finite length, N in time and a cut-off frequency W, the DPSS's are a collection of N discrete-time sequences that are strictly band-limited to the digital frequency range $|f| < W$, yet highly concentrated in time to the index range $n = 0, 1, \cdots, N-1$. Such sequences can be found using an algorithm similar to the Papoulis-Gerchberg method. Note the difference of such sequences to the ones that we intend to design in this work; we do not impose any constraints on the bandwidth of the sequences in the frequency domain. Also, the original ideas presented in this work are applicable both to finite and infinite length sequences. Moreover, we focus on the concentration of the signals in the time and frequency domain using the notion of variance.

1.4. Organization

The organization of the paper is as follows. In Section 2, we see how the definitions in Tables 2 and 3 are found starting from the continuous-time definition and using the classic notion of time and frequency uncertainties. In Section 3, we formulate precisely the problem of finding maximally compact sequences. In Section 4, we first establish key properties of maximally compact sequences to narrow down the search to real positive sequences, which then allows to setup and solve the problem numerically. From the numerical formulation, we identify the spectrum of maximally compact sequences. Finally after the simulation results and conclusions, appendices contain proofs of the stated results.

2. Theory

The choice of $\Delta^2_{\omega_p}$ as an *angular variance* must be motivated. A thorough study on the subject in [3] provides many heuristic reasons. If these reasons all have their merits, we would also like to show that the definition of a frequency spread is unambiguously linked to a pseudo-differential operator, providing a simple and unique selection criterion. We will show that $\Delta^2_{\omega_p}$ corresponds to the *finite difference* operator, which is the simplest among discrete diffrentials.

The *Heisenberg uncertainty principle* is rooted in particle physics, and therefore thought in terms of *position-momentum* uncertainty rather than *time-frequency* spread as in the signal processing community. These two interpretations are equivalent for continuous signals, but the intuitions behind them are useful to make the transition from continuous to discrete time.

2.1. Uncertainty Principle for Linear Operators

Consider a Hilbert space $\mathcal{H}$ with the inner-product $\langle \cdot\, ,\, \cdot \rangle$ and the induced norm $\|x\| \overset{\text{def}}{=} \sqrt{\langle x\, ,\, x\rangle}$. Define self-adjoint linear operators $L,\ M\ :\ \mathcal{H} \mapsto \mathcal{H}$, and their "centered" counterparts. Note that the centered operators are not linear, nor self-adjoint.

$$\bar{L}x \overset{\text{def}}{=} (L - \mu_L(x))x\,, \quad \text{where} \ \ \mu_L(x) \overset{\text{def}}{=} \langle Lx\, ,\, x\rangle \,/\, \|x\|^2\,.$$

Assuming L is self-adjoint, $\mu_L(x) \in \mathbb{R}$.

For any pair of linear operators, the *commutator*

$$[L\, ,\, M] \overset{\text{def}}{=} LM - ML\,,$$

vanishes if and only if its arguments commute [14].

With these definitions, the *Heisenberg uncertainty principle* is

Theorem 1. (Schrödinger 1930 [15])

$$\left\|\bar{L}x\right\|^2\ \left\|\overline{M}x\right\|^2 \ \geq\ \frac{1}{4}\left|\langle [L\, ,\, M]\, x\, ,\, x\rangle\right|^2\ . \tag{7}$$

Proof. Using Cauchy-Schwarz inequality and the self-adjointness of L and M,

$$\begin{aligned}
\left\|\bar{L}x\right\|^2 \left\|\overline{M}x\right\|^2 &\geq \left|\left\langle \overline{M}x\, ,\, \bar{L}x\right\rangle\right|^2 \\
&= \left|\langle LMx\, ,\, x\rangle - \mu_L(x)\,\mu_M(x)\,\|x\|^2\right|^2 \\
&= \left|\frac{1}{2}\,\langle (LM - ML)x\, ,\, x\rangle + \frac{1}{2}\,\langle (LM + ML)x\, ,\, x\rangle - \mu_L(x)\,\mu_M(x)\,\|x\|^2\right|^2\ .
\end{aligned}$$

The two halves within the modulus are respectively imaginary and real, therefore with the *anticommutator* $\{L\, ,\, M\} \overset{\text{def}}{=} LM + ML$, one obtains the *Schrödinger uncertainty principle*

$$\left\|\bar{L}x\right\|^2\ \left\|\overline{M}x\right\|^2 \ \geq\ \frac{1}{4}\left|\langle [L\, ,\, M]\, x\, ,\, x\rangle\right|^2 + \frac{1}{4}\left|\langle \{L\, ,\, M\}\, x\, ,\, x\rangle - 2\mu_L(x)\,\mu_M(x)\,\|x\|^2\right|^2\ ,$$

from which the weaker *Heisenberg uncertainty principle* immediatly follows. □

2.2. The journey from continuous to discrete

We apply the general uncertainty relation (7) to continuous and discrete-time signals to make some connections explicit. In particular, we study two schemes to make the transition from continuous-time to discrete-time.

2.2.1. Continuous time-frequency uncertainty

Let $x(t) \in L_2(\mathbb{R})$, the localization L and momentum M operators are

$$Lx(t) \stackrel{\text{def}}{=} t \cdot x(t) \ ,$$
$$Mx(t) \stackrel{\text{def}}{=} \frac{dx}{dt}(t) \ .$$

Moreover, because of the similarity between the continuous-time domain and the CTFT domain, localization and momentum are Fourier dual of each other :

$$Lx(t) \stackrel{\text{CTFT}}{\longleftrightarrow} j\frac{dX}{d\omega}(\omega) \stackrel{\text{def}}{=} M_{\mathcal{F}}X(\omega) \ ,$$
$$Mx(t) \stackrel{\text{CTFT}}{\longleftrightarrow} j\omega X(\omega) \stackrel{\text{def}}{=} L_{\mathcal{F}}X(\omega) \ .$$

This duality shows that the *localization-momentum* and *time-frequency localization* interpretations are the two sides of a same coin.

With the evaluation of

$$|\langle [t, d/dt]x \ , \ x\rangle|^2 \ = \ \|x\|^4 \ ,$$

and use of Parseval's equality

$$\left\|\overline{L_{\mathcal{F}}}X\right\|^2 / \, \|X\|^2 = \left\|\overline{M}x\right\|^2 / \, \|x\|^2 \ ,$$

the uncertainty relation (7) yields the continuous-time *Heisenberg uncertainty principle* (1)

$$\eta_c = \Delta_t^2 \ \Delta_{\omega_c}^2 \geq \frac{1}{4} \ ,$$

where $\Delta_t^2 = \left\|\bar{L}x\right\|^2 \Big/ \, \|x\|^2$ and $\Delta_{\omega_c}^2 = \left\|\overline{L_{\mathcal{F}}}X\right\|^2 \Big/ \, \|X\|^2$ as found in Table 1.

2.2.2. First attempt at discrete-time uncertainty: the DSP point of view

In the same way that we defined the localization and momentum in time and frequency for continuous-time signals, we define them for sequences. The process can be seen either from the time-domain as a discretization or from the frequency domain as a periodization.

If we follow the time-frequency interpretation, it is natural to discretize L and periodize $L_{\mathcal{F}}$

$$\begin{array}{ccc} \text{Designed Localization:} & & \text{Implied Momentum:} \\ \begin{array}{rcl} Lx_n & \stackrel{\text{def}}{=} & n \cdot x_n \\ L_{\mathcal{F}}X(e^{j\omega}) & \stackrel{\text{def}}{=} & \widetilde{\omega} \cdot X(e^{j\omega}) \end{array} & \stackrel{\text{DTFT}}{\longleftrightarrow} & \begin{array}{rcl} j\frac{dX}{d\omega}(e^{j\omega}) & \stackrel{\text{def}}{=} & M_{\mathcal{F}}X(e^{j\omega}) \\ \frac{(-1)^n}{jn} * x_n & \stackrel{\text{def}}{=} & Mx_n \end{array} \end{array} \tag{8}$$

where $\widetilde{\omega}$ represents the sawtooth wave with period 2π. The implied frequency-domain momentum $M_{\mathcal{F}}$ has the properties one would expect for momentum as it measures the variation of the spectrum locally by differentiation; on the other side, the implied time-domain momentum M does not lead to an easy interpretation.

This choice of operators not only causes interpretation problems. The uncertainty relation is

$$\eta_\ell = \Delta_n^2 \ \Delta_{\omega_\ell}^2 \geq \frac{1}{4} \left| 1 - \frac{\left|X(e^{j\pi})\right|^2 - 2\text{Re}\left\{X'(e^{j\pi})X^*(e^{j\pi})\right\}}{\|X\|^2} \right|^2 \ . \tag{9}$$

where $\Delta_n^2 \stackrel{\text{def}}{=} \|Lx\|^2 / \, \|x\|^2$ and $\Delta_{\omega_\ell}^2 \stackrel{\text{def}}{=} \|L_{\mathcal{F}}X\|^2 / \, \|X\|^2$, as found in Table 2.

Under the condition $X(e^{j\pi}) = 0$, the uncertainty principle (3) is met

$$\eta_\ell = \Delta_n^2 \ \Delta_{\omega_\ell}^2 \geq \frac{1}{4} \ , \quad \text{for } X(e^{j\pi}) = 0 \ . \tag{10}$$

However the necessity of a root at π severely reduces the utility of this result, exluding sequences such as the discretized gaussian kernel.

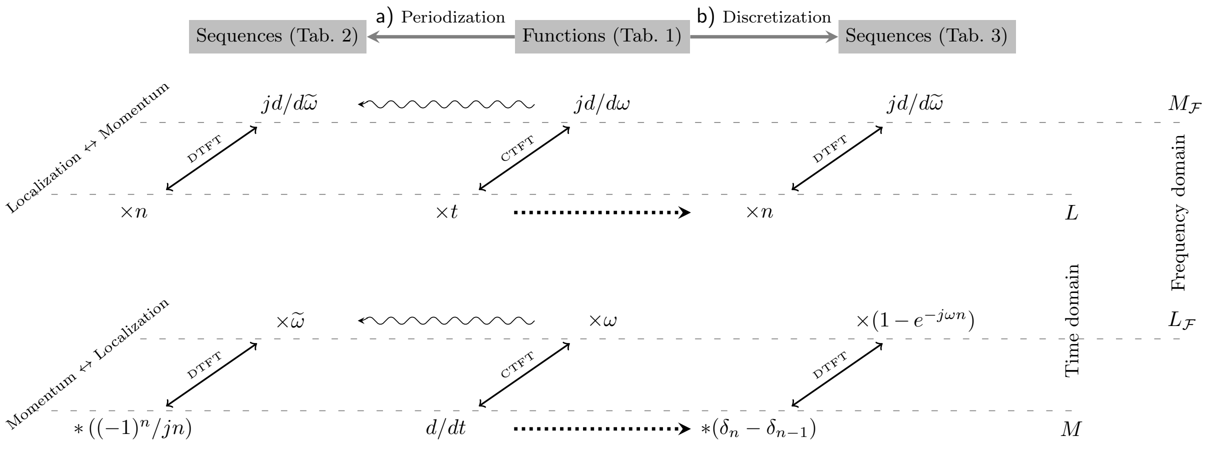

Figure 2: **Two ways to obtain an uncertainty principle for sequences**. On the left side a)—by periodization—one obtains the definitions of Table 2, on the right side b)—by discretization—one obtains the definitions of Table 3. Note that periodization and discretization yield the same result for the top plane, but not for the bottom plane.

2.2.3. Second attempt at discrete-time uncertainty: the physicist point of view

Failure of the first attempt can be pinned down on the definition of the frequency-domain localization. Instead of designing both localization operators, we may design both time-domain operators. The task is thus to discretize localization (already done) and momentum. For the latter, the *finite difference* filter is a natural candidate

$$\begin{array}{ccccccc} & \text{Designed in the time-domain} & & & & \text{Implied in the frequency-domain} & \\ Lx_n & \stackrel{\text{def}}{=} & n \cdot x_n & \stackrel{\text{DTFT}}{\longleftrightarrow} & j\frac{dX}{d\omega}(e^{j\omega}) & \stackrel{\text{def}}{=} & M_{\mathcal{F}}X(e^{j\omega}) \\ Mx_n & \stackrel{\text{def}}{=} & x_n - x_{n-1} & & (1-e^{-j\omega})\cdot X(e^{j\omega}) & \stackrel{\text{def}}{=} & L_{\mathcal{F}}X(e^{j\omega}) \end{array} \tag{11}$$

The implied operator $L_{\mathcal{F}}$ has intuitive properties. For example,

$$\|L_{\mathcal{F}}X\|^2 = 2\int_{-\pi}^{\pi}(1-\cos(\omega))\left|X(e^{j\omega})\right|^2 d\omega \ ,$$

shows that the spread is measured with respect to the smooth kernel $2(1-\cos(\omega))$ instead of $\widetilde{\omega}^2$ in (8).

Note that $2(1-\cos(\omega)) = \omega^2 + \mathcal{O}(\omega^4)$, so the two definitions of spread coincide for $\omega \to 0$, and should yield asymptotically equal results for sequences with a narrow spectrum.

The operators $M_{\mathcal{F}}$ and $L_{\mathcal{F}}$ are the ones most often used in directional statistics [4].

The right-hand side of (7) evaluates to

$$\left|\left\langle\left[1-e^{-j\omega}\ ,\ jd/d\omega\right]X\ ,\ X\right\rangle\right|^2 \ = \ \|X\|^4\,|\tau(x)|^2\ . \tag{12}$$

For $|\tau(x)| \neq 0$—which is equivalent to $\|x\|_0 > 1$—the corresponding uncertainty relation [3] is therefore

$$\eta_p = \Delta_n^2\ \Delta_{\omega_p}^2 \geq \frac{1}{4}\ , \quad \text{for } \|x\|_0 > 1, \tag{13}$$

where $\Delta_{\omega_p}^2 \stackrel{\text{def}}{=} \frac{\|\overline{L_{\mathcal{F}}}X\|^2}{|\tau(X)|^2\|X\|^2}$ and Δ_n^2 are as found in Table 3.

The study of continuous and discrete uncertainty principle shows that the definitions of time and frequency spreads found in Table 3 are not arbitrary and follow from the most natural discretization of the continuous-time definitions found in Table 1.

3. Problem Statement

As mentioned in Section 1.2, the main objective of this paper is to design maximally compact sequences as solutions of (6). Thus we are interested in solving

$$\begin{aligned} \Delta^2_{n,\mathrm{opt}} \; = \; & \underset{x_n}{\text{minimize}} \quad \Delta^2_n \\ & \text{subject to} \quad \Delta^2_{\omega_p} \; = \sigma^2 \, . \end{aligned} \tag{14}$$

where σ^2 is the fixed, given frequency spread of the sequence. We saw in (13) that the time-frequency spread of such sequences is naturally bounded from below by the Heisenberg uncertainty bound. Prestin et al. in [5] show that the lower bound in (13) is achievable only asymptotically when the frequency spread of the sequence tends to zero. Thus, the question is "what is the minimal achievable uncertainty for sequences with a given spread?". The answer to this question lies in the solution of (14).

Although finding a tight bound for maximally compact sequences is an important task, the more interesting problem is to constructively design sequences which achieve such bound. The ideal answer to this question would be to find analytic closed form solutions—like Gaussian signals for the continuous time problem—achieving the tight bound. In the following we will address these problems.

4. Main Results

We start with some properties of maximally compact sequences. These properties will greatly facilitate the task of solving (14).

4.1. Properties of Maximally Compact Sequences

In the definitions of time and frequency spreads in Table 3 we considered complex sequences and their DTFTs. In the following, we establish two lemmas that make the search for maximally compact sequences easier.

Lemma 1. Maximally compact sequences are *generalized linear-phase sequences* derived from real-valued, positive maximally compact sequences; i.e. x_n is a maximally compact sequence only if

$$\exists\, \varphi, \psi \in [0,\ 2\pi[\quad \text{such that} \quad x_n = |x_n| e^{-j(\varphi n + \psi)} \, , \tag{15}$$

where $|x|$ is a maximally compact sequence.

Proof. See Appendix A.1. □

Consider also the shift operator

$$x_{n+\nu} \quad \stackrel{\mathrm{DTFT}}{\longleftrightarrow} \quad e^{j\omega\nu} X(e^{j\omega}) \; , \quad \nu \in \mathbb{R}, \tag{16}$$

whose principal effect is to shift the time centre of a sequence

$$\mu_n(x_{n+\nu}) \; = \; \mu_n(x) - \nu. \tag{17}$$

Notice that ν might be non-integer, in which case $x_{n-\nu}$ is a shorthand for sinc resampling on a grid shifted by ν in the time-domain.

Lemma 2. If x is a maximally compact sequence, then $x_{n-\mu_n(x)}$ is also maximally compact.

Proof. See Appendix A.2. □

Lemmas 1 and 2 greatly reduce the complexity of the problem, and from here on we only consider—without loss of generality—real, positive sequences x, with $\mu_n(x) = 0$ and $\|x\|^2 = 1$.

4.2. Main Theorems

The following theorem is the core of the results presented in this paper.

Theorem 2. For finding the solution of (14), it is sufficient to solve the following semi-definite program (SDP)

$$\begin{aligned} \underset{\boldsymbol{X}}{\text{minimize}} \quad & \operatorname{tr}(\boldsymbol{A}\boldsymbol{X}) \\ \text{subject to} \quad & \operatorname{tr}(\boldsymbol{B}\boldsymbol{X}) = \alpha \\ & \operatorname{tr}(\boldsymbol{X}) = 1, \quad \boldsymbol{X} \succeq 0\,, \end{aligned} \tag{18}$$

where $\alpha = \frac{1}{\sqrt{1+\sigma^2}}$. Further, $\boldsymbol{X}^{\mathrm{opt}}$, the solution to (18) has rank one and $\boldsymbol{X}^{\mathrm{opt}} = \boldsymbol{x}^{\mathrm{opt}}\,\boldsymbol{x}^{\mathrm{opt}T}$, with $\boldsymbol{x}^{\mathrm{opt}}$ the solution of (14). Matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are defined as

$$\boldsymbol{A} = \begin{bmatrix} \ddots & & & & & \\ & 2^2 & & & 0 & \\ & & 1^2 & & & \\ & & & 0 & & \\ & & & & 1^2 & \\ & 0 & & & & 2^2 \\ & & & & & & \ddots \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \ddots & & & & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ & \frac{1}{2} & 0 & \frac{1}{2} & \\ & & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & & & & \ddots \end{bmatrix}. \tag{19}$$

Proof. See Appendix B. $\square$

The SDP in (18) can be solved to an arbitrary precision by using existing approaches in the optimization literature; for example using the cvx software package [16]. This, gives a constructive way to design sequences that are maximally compact in the time domain with a given frequency spread.

Example 2. Take $\sigma^2 = 0.1$ to be the fixed and given frequency spread of the sequence. We can use cvx to solve the semi-definite program (18) and find the optimal value of $\Delta_n^2 = 2.62$. This results in the time-frequency spread of $\eta_p = 0.262$. The simple code in MATLAB is:

```
cvx_begin
variable X(n,n);
minimize(trace(A*X))
subject to
    trace(B*X) == 1/sqrt(1+0.1)
    trace(X) == 1
    X == semi-definite(n)
cvx_end;
```

Note that contrary to continuous-time signals, we cannot reach the 0.25 lower bound for sequences. The resulting sequence and its DTFT are shown in Figure 3.

The dual of SDP (18) is [17]:

$$\begin{aligned} \underset{\lambda_1,\lambda_2}{\text{maximize}} \quad & \alpha\,\lambda_1 + \lambda_2 \\ \text{subject to} \quad & \boldsymbol{A} - \lambda_1\,\boldsymbol{B} - \lambda_2\,\boldsymbol{I} \succeq 0 \end{aligned} \tag{20}$$

Lemma 3. For the primal problem (18) and the dual (20), strong duality holds.

Proof. We refer the reader to Appendix C for the proof. $\square$

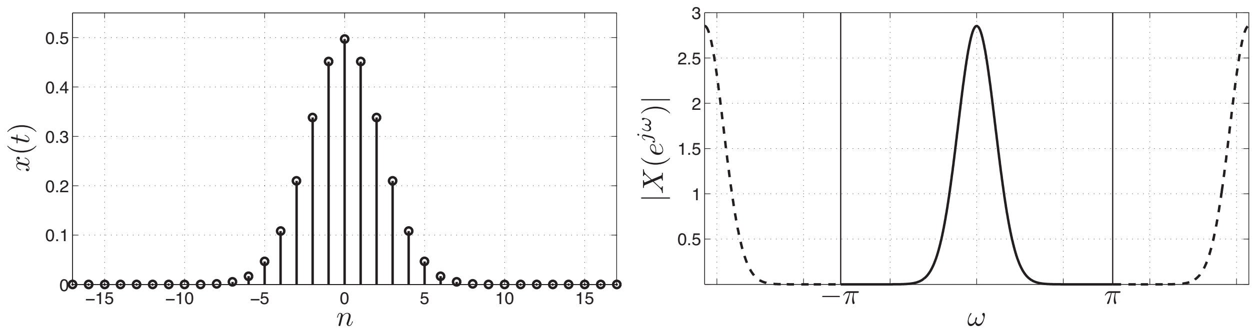

Figure 3: **An example solution of** (18)**.** The output of the SDP in (18) with $\sigma^2 = 0.1$ using cvx in Example 2. The optimal value for Δ_n^2 is found to be 2.62 which results in a time-frequency spread of $\eta_p = 0.262$.

Thus, for finding the time-frequency spread of maximally compact sequences, solving the dual problem suffices.

Note that although Theorem 2 provides a constructive way for finding maximally compact sequences, it does not specify the closed form for these sequences. One would be interested to see if—in analogy to continuous-time—sampled gaussians are maximally compact? The answer is negative, as shown by this theorem:

Theorem 3. The DTFT spectrum, $X(e^{j\omega})$ of maximally compact sequences are Mathieu functions. More specifically,

$$X(e^{j\omega}) \;=\; \gamma_0 \cdot \mathrm{ce}_0(-2\lambda_1\,;\,(\omega-\omega_0)/2)e^{j\mu\omega}\ , \tag{21}$$

where $|\gamma_0| = \|\mathrm{ce}_0(-2\lambda_1\,;\,(\omega-\omega_0)/2)\|^{-1}$, ω_0 and μ are shifts in frequency or time and λ_1 is the optimal solution of the dual problem (20). $\mathrm{ce}_0(q\,;\omega)$ is Mathieu's harmonic cosine function of order zero.

For the proof of the theorem and further insights about Mathieu functions, we refer the reader to Appendix D.

Using the constructive method presented in Theorem 2, we can find the achievable (and tight) uncertainty principle bound for discrete sequences. This is shown and discussed more in Section 5 and Figure 4. However, a numerically computed boundary may not always be practical, and even though the numerical solution exactly solves the problem, its accuracy may be challenged. Therefore, we characterize the asymptotic behavior of the time-frequency bound:

Theorem 4. If x_n is maximally compact for a given $\Delta_{\omega_p}^2 = \sigma^2$, then

$$\eta_p = \Delta_n^2\ \Delta_{\omega_p}^2 \geq \sigma^2\left(1-\sqrt{\frac{\sigma^2}{1+\sigma^2}}\right). \tag{22}$$

Proof. The proof for this theorem is provided in Appendix E. □

This fundamental result states that for a given frequency spread, we cannot design sequences which achieve the classic Heisenberg uncertainty bound. We will see how this curve compares to the classic Heisenberg bound in Section 5.

The lower bound in (22) converges to 1/2 as the value of σ^2 grows, and "pushes up" the time-frequency spread of maximally compact sequences towards 1/2 which is also an asymptotic upper bound on the time-frequency spread as $\Delta_{\omega_p}^2 \to \infty$; indeed, one may construct the unit-norm sequence $x_n^{(\varepsilon)} = \varepsilon\delta_{n+1} + \sqrt{1-2\varepsilon^2}\delta_n + \varepsilon\delta_{n-1}$, which verifies $\lim_{\varepsilon\to 0}\eta_p(x^{(\varepsilon)}) = 1/2$.

Theorem 5. For small values of σ^2, maximally compact sequences satisfy

$$\eta_p = \Delta_n^2\ \Delta_{\omega_p}^2 \leq \frac{\sigma^2}{8}\left(\frac{\sqrt{1+\sigma^2}}{\sqrt{1+\sigma^2}-1}-\frac{1}{2}\right). \tag{23}$$

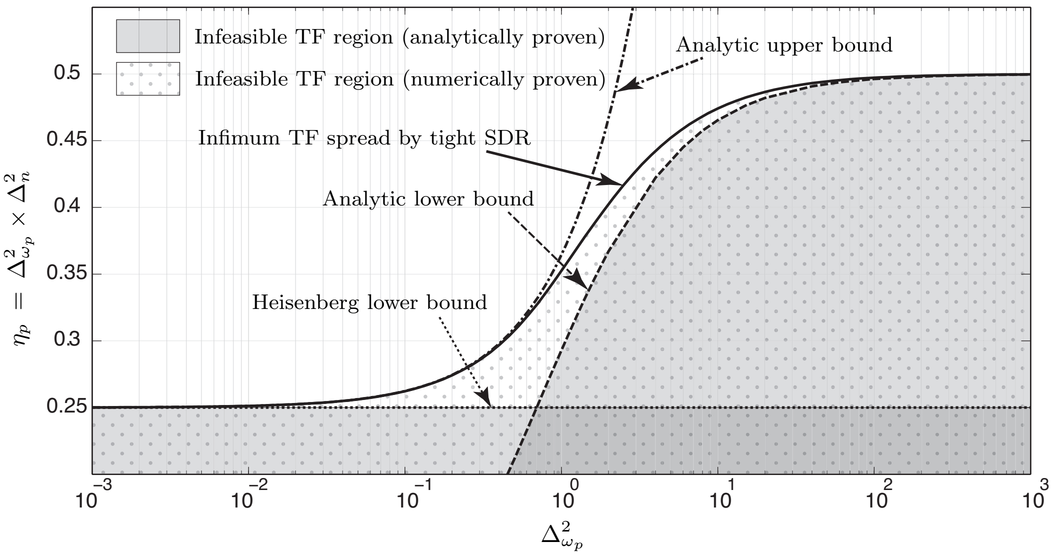

Figure 4: **New uncertainty bounds**. The solid line shows the results of solving the SDP in (18). The dotted line shows the classic Heisenberg uncertainty principle. The dashed lines show the analytic lower and upper bounds found in Theorems 4 and 5, respectively.

Proof. See Appendix F. □

For small values of σ^2, the upper bound in (23) converges from above to 1/4, thus "pushing down" the time-frequency spread of maximally compact sequences towards the Heisenberg uncertainty bound 1/4 from above.

4.2.1. Finite-Length Sequences

The theory that we have provided so far holds for infinite sequences. For computational purposes, we have to assume finite length for the sequences in the time domain, which is not an issue if the sequence length is chosen to be long enough. As a side benefit, a length constraint on the sequence may be put at will without changing the design algorithm.

5. Simulation Results

In order to show the behaviour of the results obtained in Theorems 2, 4 and 5, we ran some simulations. For this end, we assumed that the designed filter is finite length with 201 taps in the time domain (the length is long enough not to pose restrictions on the solution). For different values of $\Delta^2_{\omega_p} = \sigma^2$, we solved the semi-definite program (18) using the cvx toolbox in MATLAB.

The resulting values of Δ^2_n were then multiplied with the corresponding $\Delta^2_{\omega_p}$ to produce the time-frequency spread of maximally compact sequences. The time-frequency spread of maximally compact sequences versus their frequency spread is shown with the solid curve in Figure 4. This means—numerically—that any time-frequency spread under this curve is not achievable. The dotted line in this figure shows the classic Heisenberg uncertainty bound. Comparing the two curves shows the gap between the classic Heisenberg principle and what is achievable in practice. The dashed lines represent analytical lower and upper bounds for the time-frequency spread of maximally compact sequences (found in Theorems 4 and 5, respectively).

Further, to give an insight on how the time-frequency spread of some common filters compare to that of maximally compact sequences, we plot their time-frequency spread together with the new uncertainty bound in Figure 5. By changing the length of each filter in time, we can find its time and frequency spreads which results in a point on the figure. We observe that as shown by Prestin et al. in [6], asymptotically

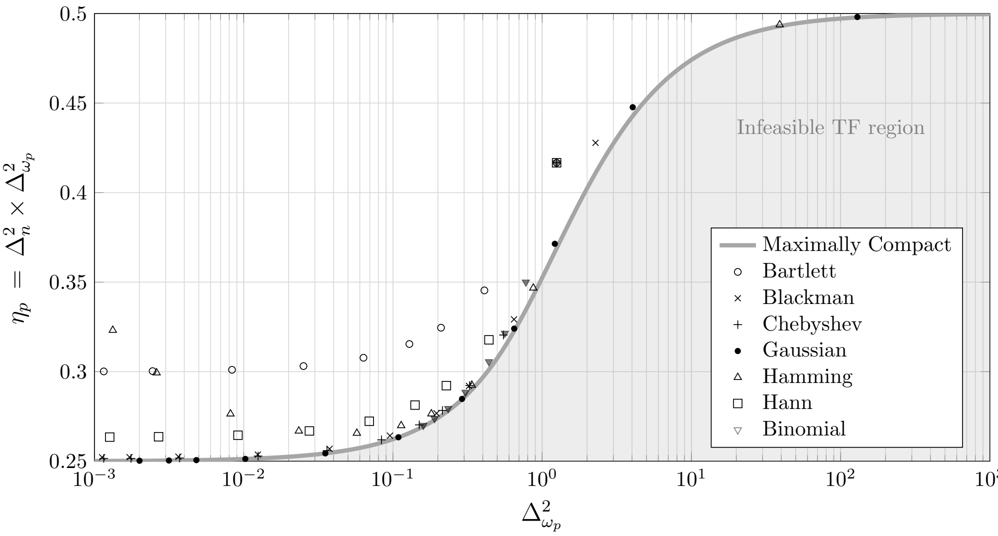

Figure 5: **Time-frequency spread of common FIR filters.** By changing the length of the filters in time, we compute the time and frequency spreads for each type of the filters. For small values of frequency spread, Gaussian filters are good approximations of Mathieu functions (as shown also in [6]).

when the frequency spread of sequences are very small, sampled Gaussians converge to the lower bound for maximally compact sequences.

6. Conclusion

We showed that contrary to continuous-time signals, for discrete sequences, there is a large gap between the Heisenberg uncertainty bound and the time-frequency spread of maximally compact sequences. Note that our results are consistent with the conclusions in [6] that no sequence can achieve the 1/4 Heisenberg uncertainty bound. The constructive way that we propose in this research enables us to design sequences that are closest to the Heisenberg bound. Moreover, having Mathieu functions as the sprectra of maximally compact sequences, shows that again in contrary to continuous-time signals, sampled Gaussians are not the minimizers of the time-frequency spread, though they are close.

Ackowledgements

The authors would like to thank Dr. Arash Amini for his help in proving Theorem 4, and Ivan Dokmanic for his insight on Mathieu functions.

Appendix A. Proof of Lemmas 1 and 2

Appendix A.1. Proof of Lemma 1

Let x be maximally compact for a given time-spread $\Delta_n^2(x)$.

$$\Delta_n^2(x) \;=\; \Delta_n^2(|x|).$$

$$\begin{aligned} \Delta_{\omega_p}^2(|x|) \;&=\; \left|\sum_{n\in\mathbb{Z}} |x_n||x_{n+1}|\right|^{-2} - 1, \\ &\leq \left|\sum_{n\in\mathbb{Z}} x_n x_{n+1}^*\right|^{-2} - 1 \quad = \Delta_{\omega_p}^2(x). \end{aligned} \tag{A.1}$$

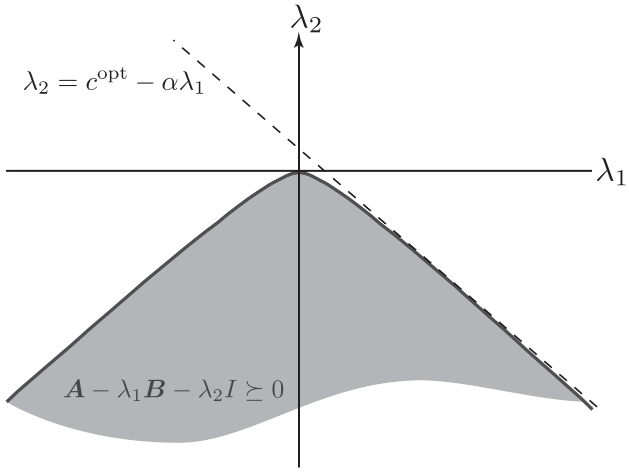

Figure A.6: **The feasible set of the dual problem** (20) **and the supporting line**. As α increases, we need to elevate the line more to support the feasible set, which means that the optimal value of Δ_n^2 increases.

Since x is maximally compact, equality must be met in (A.1) to avoid contradiction, which amounts to

$$x_n x_{n+1}^* \; = \; |x_n||x_{n+1}|e^{j\varphi}, \; \varphi \in [0, \; 2\pi) \, , \; \forall n \in \mathbb{Z}.$$

This condition is equivalent to (15).

For maximally compact sequences, if Δ_n^2 strictly monotonically varies in function of $\Delta_{\omega_p}^2$, then fixing Δ_n^2 or $\Delta_{\omega_p}^2$ is equivalent and proves the lemma. In the following lemma we show that for maximally compact sequences Δ_n^2 changes monotonically with $\Delta_{\omega_p}^2$.

Lemma 4. For maximally compact sequences, Δ_n^2 is a decreasing function of $\Delta_{\omega_p}^2$.

Proof. For proving this lemma, we use the dual formulation in (20). The feasible region of the dual problem is shown in Figure A.6. We can write the dual as

$$\begin{aligned} \underset{\lambda_1,\lambda_2}{\text{maximize}} \quad & c \\ \text{subject to} \quad & \lambda_2 = c - \alpha\,\lambda_1\,, \\ & \boldsymbol{A} - \lambda_1\,\boldsymbol{B} - \lambda_2\,\boldsymbol{I} \succeq 0\,. \end{aligned} \tag{A.2}$$

Note that α changes between 0 and 1 (see Figure A.6). For a fixed α, the maximum c^{opt} is found by elevating the corresponding line $\lambda_2 = c - \alpha\lambda_1$ until it supports the feasible set (it is tangent to it). Since the feasible set is convex, as α grows (which means $\Delta_{\omega_p}^2$ decreases), we need a higher elevation of the line to support the convex set, thus c^{opt} (equivalently Δ_n^2) increases. □

Appendix A.2. Proof of Lemma 2

Consider the shift operator in (16). Since the shift operation does not change the norm of a sequence, we will assume a unit norm sequence without loss of generality. We can show that

$$\begin{aligned} \mu_n(x_{n-\nu}) \; &= \frac{\left\langle j\frac{d}{d\omega}\left(e^{-j\omega\nu}X(e^{j\omega})\right), e^{-j\omega\nu}X(e^{j\omega})\right\rangle}{2\pi} \\ &= \frac{\left\langle j\frac{d}{d\omega}X(e^{j\omega}), X(e^{j\omega})\right\rangle}{2\pi} + \frac{\nu\left\langle X(e^{j\omega}), X(e^{j\omega})\right\rangle}{2\pi} \\ &= \mu_n(x) + \nu, \end{aligned} \tag{A.3}$$

where we used the DTFT domain definition of the time center [6]:

$$\mu_n(x) \;=\; \frac{\left\langle j\frac{d}{d\omega}X(e^{j\omega}), X(e^{j\omega})\right\rangle}{\|X\|^2}.$$

The proof for Lemma 2—trivial for $\nu \in \mathbb{Z}$—is not obvious for arbitrary shifts. Let x be maximally compact with time center $\mu_n(x)$, then according to (A.3), $x_{n-\nu}$ is centered at $\mu_n(x)+\nu$ and

$$\begin{aligned}
2\pi\Delta_n^2(x_{n-\nu}) &= 2\pi\left[\sum_{n\in\mathbb{Z}} n^2(x_{n-\nu})^2 - |\mu_n(x_{n-\nu})|^2\right] \\
&= \left\langle j\frac{d}{d\omega}e^{-j\omega\nu}X(e^{j\omega}), j\frac{d}{d\omega}e^{-j\omega\nu}X(e^{j\omega})\right\rangle - 2\pi\,|\mu_n(x)+\nu|^2 \\
&= \left\langle -j\nu X(e^{j\omega}) + X'(e^{j\omega}), -j\nu X(e^{j\omega}) + X'(e^{j\omega})\right\rangle - 2\pi\,|\mu_n(x)+\nu|^2 \\
&= 2\pi\,|\nu|^2\left\langle X(e^{j\omega}), X(e^{j\omega})\right\rangle + \left\langle X'(e^{j\omega}), X'(e^{j\omega})\right\rangle \\
&\qquad + \left\langle -j\nu X(e^{j\omega}), X'(e^{j\omega})\right\rangle + \left\langle X'(e^{j\omega}), -j\nu X(e^{j\omega})\right\rangle - 2\pi\,|\mu_n(x)+\nu|^2 \\
&= 2\pi\,|\nu|^2 + \left\langle X'(e^{j\omega}), X'(e^{j\omega})\right\rangle + 2\pi\ \mathrm{Real}[\mu_n(x)\nu^*] - 2\pi\,|\mu_n(x)+\nu|^2 \\
&= \left\langle X'(e^{j\omega}), X'(e^{j\omega})\right\rangle - 2\pi\,|\mu_n(x)|^2 \\
&= 2\pi\Delta_n^2(x)\,.
\end{aligned} \tag{A.4}$$

This shows that time shift does not affect the time spread of a sequence. Thus, if x is a maximally compact sequence, then $x_{n-\mu_n(x)}$ is also maximally compact (note that time shift does not change the frequency characteristics of the sequence). □

Appendix B. Proof of Theorem 2

By using Lemmas 1 and 2, we can write problem (14) as

$$\begin{aligned}
\Delta^2_{n,\mathrm{opt}} \;=\; &\underset{x_n}{\text{minimize}} && \sum_{n\in\mathbb{Z}} n^2 x_n^2 \\
&\text{subject to} && \sum_{n\in\mathbb{Z}} x_n x_{n+1} = \frac{1}{\sqrt{1+\sigma^2}}, \\
& && \sum_{n\in\mathbb{Z}} x_n^2 = 1.
\end{aligned} \tag{B.1}$$

We can rewrite (B.1) in a matrix form as a quadratically constrained quadratic program (QCQP) [17]

$$\begin{aligned}
&\underset{\boldsymbol{x}}{\text{minimize}} && \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} \\
&\text{subject to} && \boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} = \alpha\,, \\
& && \boldsymbol{x}^T\boldsymbol{x} = 1\,,
\end{aligned} \tag{B.2}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are defined in (19) and $\alpha = 1/\sqrt{1+\sigma^2}$. This problem can be further reformulated as follows:

$$\begin{aligned}
&\underset{\boldsymbol{x}}{\text{minimize}} && \mathrm{tr}(\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^T) \\
&\text{subject to} && \mathrm{tr}(\boldsymbol{B}\boldsymbol{x}\boldsymbol{x}^T) = \alpha \\
& && \mathrm{tr}(\boldsymbol{x}\boldsymbol{x}^T) = 1\,.
\end{aligned}$$

Replacing $\boldsymbol{x}\boldsymbol{x}^T$ by $\boldsymbol{X}$, we can write equivalently

$$
\begin{aligned}
\underset{\boldsymbol{X}}{\text{minimize}} \quad & \operatorname{tr}(\boldsymbol{A}\boldsymbol{X}) \\
\text{subject to} \quad & \operatorname{tr}(\boldsymbol{B}\boldsymbol{X}) = \alpha \\
& \operatorname{tr}(\boldsymbol{X}) = 1 \\
& \boldsymbol{X} \succeq 0,\ \operatorname{rank}(\boldsymbol{X}) = 1\,.
\end{aligned}
\tag{B.3}
$$

We further relax the above formulation to reach the semi-definite program

$$
\begin{aligned}
\underset{\boldsymbol{X}}{\text{minimize}} \quad & \operatorname{tr}(\boldsymbol{A}\boldsymbol{X}) \\
\text{subject to} \quad & \operatorname{tr}(\boldsymbol{B}\boldsymbol{X}) = \alpha \\
& \operatorname{tr}(\boldsymbol{X}) = 1, \quad \boldsymbol{X} \succeq 0\,.
\end{aligned}
\tag{B.4}
$$

In Lemma 5 we show that the semi-definite relaxation is tight.

□

Lemma 5. The semi-definite relaxation (SDR) from (B.3) to (B.4) is tight.

Proof. Shapiro and then Barnivok and Pataki [18, 19, 20, 21] show that if the SDP in (B.3) is feasible, then

$$
\operatorname{rank}(\boldsymbol{X}^{\text{opt}}) \leq \lfloor (\sqrt{8m+1} - 1)/2 \rfloor\,, \tag{B.5}
$$

where m is the number of constraints of the SDP and $\boldsymbol{X}^{\text{opt}}$ is its optimal solution. For our semi-definite program in (B.3), $m = 2$. Thus, (B.5) implies that the solution has rank 1. Using this fact, one can see that the semi-definite relaxation is in fact tight. Note that from the nature of the problem, (B.3) is clearly feasible; we can always find a periodic signal in the Fourier domain with a unit norm and a desired frequency spread, although not having an optimal time spread. □

Appendix C. Proof of Lemma 3

We use the following Lemma for the proof:

Lemma 6. [22, 23] For a semi-definite program and its dual: If the primal is feasible and the dual is strictly feasible, then strong duality holds.

It is easy to see the primal is feasible. One can always find a signal $\boldsymbol{x}$ with a certain frequency spread and norm one. Using this signal we can construct $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$ which shows the feasibility of the primal. For the dual, one can use the Gershgorin's circle theorem and show that a sufficient condition for $\boldsymbol{A} - \lambda_1 \boldsymbol{B} - \lambda_2 \boldsymbol{I} \succ 0$ to hold is $\lambda_2 < -\lambda_1$ and $\lambda_1 > 0$. Thus, the dual problem is strictly feasible. □

Appendix D. Proof of Theorem 3

If a sequence is a solution to the dual SDP problem (20), the dual constraint is active. Therefore, maximally compact sequences lie on the boundary of the quadratic cone

$$
\boldsymbol{A} - \lambda_1 \boldsymbol{B} - \lambda_2 \boldsymbol{I} \succeq 0.
$$

A maximally compact sequence $\boldsymbol{x}$ is thus solution of the eigenvalue problem

$$
(\boldsymbol{A} - \lambda_1 \boldsymbol{B})\boldsymbol{x} \;=\; \lambda_2 \boldsymbol{x}, \tag{D.1}
$$

where λ_1 and λ_2 are the dual variables of the SDP problem, and λ_2 is also the minimal eigenvalue of $\boldsymbol{A} - \lambda_1 \boldsymbol{B}$ and x is the associated eigenvector (this can be also seen by forcing the derivative of the Lagrangian in (B.2) to zero).

This explicit link between the dual variables and the sequence, yields a differential equation for which the DTFT spectrum of maximally compact sequences is the solution. In the DTFT domain (D.1) becomes (expanding the matrix multiplications)

$$\begin{aligned} & -X''(e^{j\omega}) - \lambda_1 \cos(\omega) X(e^{j\omega}) \;=\; \lambda_2 X(e^{j\omega}), \\ \Leftrightarrow \qquad & X''(e^{j\omega}) + (\lambda_2 + \lambda_1 \cos(\omega))\, X(e^{j\omega}) \;=\; 0\,, \end{aligned} \tag{D.2}$$

which is *Mathieu's differential equation* ([24]§20.1.1):

$$\frac{\partial^2 y(t)}{\partial t^2} + (a - 2q\cos(2t)) \cdot y(t) \;=\; 0. \tag{D.3}$$

The solutions of Mathieu's equation are called Mathieu functions, and they assume an odd and even form

$$\text{Mathieu's Cosine (even)} \quad \mathrm{ce}(a, q; \omega)\ , \tag{D.4}$$

$$\text{Mathieu's Sine (odd)} \quad \mathrm{se}(a, q; \omega)\ . \tag{D.5}$$

Taking into account the periodicity of (D.2), it appears not all pairs of parameters (a, q) will lead to a periodic solution. Mathieu functions can be restricted to be 2π periodic:

Definition 1. The solutions of Mathieu's harmonic differential equation—equation (D.3) with a solution y 2π-periodic—are defined as

$$\text{Mathieu's harmonic Cosine (even, periodic)} \quad \mathrm{ce}_m(q\,;\,\omega) \;=\; \mathrm{ce}(a_m(q), q\,;\,\omega)\ ,\ m \in \mathbb{N}. \tag{D.6}$$

$$\text{Mathieu's harmonic Sine (odd, periodic)} \quad \mathrm{se}_m(q\,;\,\omega) \;=\; \mathrm{ce}(b_m(q), q\,;\,\omega)\ ,\ m \in \mathbb{N}^+. \tag{D.7}$$

It is immediately visible that the spectrum of maximally compact sequences may only have the form

$$\begin{aligned} X(e^{j\omega}) \; &= \gamma_0 \cdot \mathrm{ce}_m(-2\lambda_1\,;\,\omega/2) + \gamma_1 \cdot \mathrm{se}_m(-2\lambda_1\,;\,\omega/2),\ \text{for } m \in \mathbb{N}^+, \\ X(e^{j\omega}) \; &= \gamma_0 \cdot \mathrm{ce}_m(-2\lambda_1\,;\,\omega/2),\ \text{for } m = 0\,, \end{aligned} \tag{D.8}$$

for any constants γ_0 and γ_1 such that $\|X(e^{j\omega}\| = 1$. More specifically, for any $\lambda_1 \geq 0$, the dual SDP problem can be posed and any solution would have the form (D.8)

Charasteristic numbers of Mathieu's equation are ordered [25], such that

$$\begin{aligned} & a_0(-2\lambda_1) < a_1(-2\lambda_1) < b_1(-2\lambda_1) < b_2(-2\lambda_1) < a_2(-2\lambda_1) < \cdots \qquad , \lambda_1 > 0\,, \\ & a_0(-2\lambda_1) < b_1(-2\lambda_1) \leq a_1(-2\lambda_1) < b_2(-2\lambda_1) \leq a_2(-2\lambda_1) < \cdots \qquad , \lambda_1 \leq 0\,, \end{aligned}$$

By (D.3) and with the substitution $t \to \omega/2$ one obtains $a_m(-2\lambda_1) = 4\lambda_2$. Because λ_2 is the minimal eigenvalue, we conclude that $m = 0$.

Note that this result validates the one in [26] which stated that asymptotically Mathieu functions minimize the time-frequency product. With a diferent approach, we can establish that only Mathieu's harmonic cosine of order 0 minimizes this product for any given frequency-spread. $\square$

Appendix E. Proof of Theorem 4

In order to prove the theorem we first provide the following lemma.

Lemma 7. If

$$\lambda_2 \leq 1 - \sqrt{1 + \lambda_1^2}\,, \tag{E.1}$$

then $\boldsymbol{P} = \boldsymbol{A} - \lambda_1 \boldsymbol{B} - \lambda_2 \boldsymbol{I} \succ 0$.

Proof. Note that if $\lambda_1 = 0$, then the matrix $\boldsymbol{P}$ is positive-semi-definite with the given condition. So in the proof we will assume that $\lambda_1 \neq 0$.

Consider the following matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$:

$$\boldsymbol{P}_1 = \begin{bmatrix} 0-\lambda_2 & -\lambda_1 & 0 & 0 & 0 \\ -\lambda_1/2 & 1-\lambda_2 & -\lambda_1/2 & 0 & 0 \\ 0 & -\lambda_1/2 & 4-\lambda_2 & -\lambda_1/2 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots \end{bmatrix}, \boldsymbol{P}_2 = \begin{bmatrix} 1-\lambda_2 & -\lambda_1/2 & 0 & 0 \\ -\lambda_1/2 & 4-\lambda_2 & -\lambda_1/2 & 0 \\ 0 & \ddots & \ddots & \ddots \end{bmatrix} \tag{E.2}$$

Call $\mathcal{I}$ the set of eigenvalues of $\boldsymbol{P}$, $\mathcal{I}_1$ the set of eigenvalues of $\boldsymbol{P}_1$ and $\mathcal{I}_2$ the set of eigenvalues of $\boldsymbol{P}_2$. It is trivial to see that $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$.

We show that if condition (E.1) is satisfied, then both $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ have positive eigenvalues.

1. $\boldsymbol{P}_1$
   One can decompose $\boldsymbol{P}_1$ as

$$\boldsymbol{P}_1 = \begin{bmatrix} 2 & & & 0 \\ & 1 & & \\ & & 1 & \\ & & & \ddots & \\ 0 & & & & 1 \end{bmatrix} \times \begin{bmatrix} -\lambda_2/2 & -\lambda_1/2 & 0 & 0 & 0 \\ -\lambda_1/2 & 1-\lambda_2 & -\lambda_1/2 & 0 & 0 \\ 0 & -\lambda_1/2 & 4-\lambda_2 & -\lambda_1/2 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots \end{bmatrix} = \boldsymbol{D} \times \boldsymbol{P}_1^{(\mathrm{s})} . \tag{E.3}$$

Note that both $\boldsymbol{D}$ and $\boldsymbol{P}_1^{(\mathrm{s})}$ are symmetric. Now define

$$\begin{aligned} \boldsymbol{S} &\stackrel{\text{def}}{=} \boldsymbol{D}^{-1/2}\boldsymbol{P}_1\boldsymbol{D}^{1/2} \\ &= \boldsymbol{D}^{-1/2}\left(\boldsymbol{D}\boldsymbol{P}_1^{(\mathrm{s})}\right)\boldsymbol{D}^{1/2} \\ &= \boldsymbol{D}^{1/2}\boldsymbol{P}_1^{(\mathrm{s})}\boldsymbol{D}^{1/2} \end{aligned}$$

First observe that $\boldsymbol{S}$ is symmetric, thus it has real eigenvalues.

Lemma 8. Eigenvalues of $\boldsymbol{S}$ and $\boldsymbol{P}_1$ are equal. Further, if $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{S}$, then $\boldsymbol{D}^{1/2}\boldsymbol{v}$ is an eigenvector of $\boldsymbol{P}_1$.

Proof. If λ is an eigenvalue of $\boldsymbol{S}$ corresponding to eigenvector $\boldsymbol{v}$, then

$$\begin{aligned} & \boldsymbol{S}\,\boldsymbol{v} = \lambda\,\boldsymbol{v} \\ \Rightarrow\quad & \boldsymbol{D}^{-1/2}\,\boldsymbol{P}_1\,\boldsymbol{D}^{1/2} = \lambda\,\boldsymbol{v} \\ \Rightarrow\quad & \boldsymbol{P}_1\,\boldsymbol{D}^{1/2}\,\boldsymbol{v} = \boldsymbol{D}^{1/2}\,\lambda\,\boldsymbol{v} \\ \Rightarrow\quad & \boldsymbol{P}_1\left(\boldsymbol{D}^{1/2}\,\boldsymbol{v}\right) = \lambda\left(\boldsymbol{D}^{1/2}\,\boldsymbol{v}\right) \end{aligned}$$

□

According to Lemma 8, it suffices to consider eigenvalues of $\boldsymbol{S}$; If $\boldsymbol{S}$ is positive-definite then all the eigenvalues of $\boldsymbol{P}_1$ are positive, and

$$\boldsymbol{S} = \begin{bmatrix} -\lambda_2 & -\lambda_1/\sqrt{2} & 0 & 0 & 0 \\ -\lambda_1/\sqrt{2} & 1-\lambda_2 & -\lambda_1/2 & 0 & 0 \\ 0 & -\lambda_1/2 & 4-\lambda_2 & -\lambda_1/2 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots \end{bmatrix} .$$

Sylvester's criterion states that a symmetric matrix is positive definite if and only if its principal minors are all positive. We can use Gaussian elimination on the matrix $\boldsymbol{S}$ to compute its principal minors.

$$\boldsymbol{S}^{\mathrm{U}} = \begin{bmatrix} 0-\lambda_2 & -\lambda_1/\sqrt{2} & 0 & 0 & 0 \\ 0 & s_1 & -\lambda_1/2 & 0 & 0 \\ 0 & 0 & s_2 & -\lambda_1/2 & 0 \\ 0 & 0 & 0 & \ddots & \ddots \end{bmatrix}, \tag{E.4}$$

where $s_1 = 1 - \lambda_2 + \lambda_1^2/2\lambda_2$ and $s_2 = 4 - \lambda_2 - \lambda_1^2/4s_1$. In general one can write

$$s_{i+1} = (i+1)^2 - \lambda_2 - \frac{\lambda_1^2}{4s_i}, \quad i \geq 1 .$$

Note that

$$s_{i+2} - s_{i+1} = 2i + 3 + \lambda_1^2 \frac{s_{i+1} - s_i}{4 s_i s_{i+1}} .$$

If $s_{i+1} > s_i$ then $s_{i+2} > s_{i+1}$. Thus if we have $s_0 > 0$ and $s_2 > s_1$, then by induction, $s_i > 0$ and thus $\boldsymbol{S}$ is positive-definite. We show that under condition (E.1), this is true. One can easily check that under this condition:

(i) $s_0 > 0$.

(ii) $s_1 = \frac{\sqrt{1+\lambda_1^2}-1}{2} > 0$.

(iii) $s_2 - s_1 = 3 > 0$.

This shows that under condition (E.1) $\boldsymbol{S}$ is positive definite and thus $\boldsymbol{P}_1$ has only positive eigenvalues.

2. $\boldsymbol{P}_2$:
As $\boldsymbol{P}_2$ is already symmetric, we can immediately use Sylvester's criterion on it. One can write

$$\boldsymbol{P}_2^{\mathrm{U}} = \begin{bmatrix} 1-\lambda_2 & -\lambda_1/2 & 0 & 0 & 0 \\ 0 & s_2 & -\lambda_1/2 & 0 & 0 \\ 0 & 0 & s_3 & -\lambda_1/2 & 0 \\ 0 & 0 & 0 & \ddots & \ddots \end{bmatrix}, \tag{E.5}$$

where

$$s_{i+1} = (i+1)^2 - \lambda_2 - \frac{\lambda_1^2}{4s_i}, \quad i \geq 1 .$$

It is easy to see that under condition (E.1), $s_i > |\lambda_1|/2 > 0$:

(i)
$$s_1 = 1 - \lambda_2 = 1 - (1 - \sqrt{1+\lambda_1^2}) = \sqrt{1+\lambda_1^2} > \frac{|\lambda_1|}{2} .$$

(ii)
$$s_{i+1} = (i+1)^2 - \lambda_2 - \frac{\lambda_1^2}{4s_i} = i^2 + 2i + \sqrt{1+\lambda_1^2} - \frac{\lambda_1^2}{2s_i} .$$

Thus, by induction if $s_i > |\lambda_1|/2$, then

$$\begin{aligned} s_{i+1} &> i^2 + 2i + |\lambda_1| - \frac{\lambda_1^2}{4\,|\lambda_1|/2} \\ &= i^2 + 2i + |\lambda_1| - \frac{|\lambda_1|}{2} \\ &> \frac{|\lambda_1|}{2} \\ &> 0 . \end{aligned}$$

Thus, $\boldsymbol{P}_2$ is also positive-definite. Putting these together, one can conclude that under condition (E.1), the matrix $\boldsymbol{P}$ is positive-definite.

$\square$

Note that condition (E.1) gives a sufficient (but not necessary) condition on the feasible set of the dual problem (20). Thus, it provides a lower bound for the maximum value of the dual. Consider the restricted dual problem

$$\begin{aligned} &\underset{\lambda_1,\lambda_2}{\text{maximize}} && \alpha\,\lambda_1+\lambda_2 \\ &\text{subject to} && \lambda_2 \le 1-\sqrt{1+\lambda_1^2} \end{aligned} \tag{E.6}$$

The solution to this problem is simply $1-\sqrt{1-\alpha^2}$. If we rewrite α in terms of σ^2, we have at the end

$$\Delta^2_{n,\text{opt}} \ge 1-\sqrt{\frac{\sigma^2}{1+\sigma^2}}\,.$$

This concludes the proof for the lower bound. $\square$

Appendix F. Proof of Theorem 5

It is easy to see that for small values of σ^2 (equivalently α closer to 1), the maximum of the dual problem is achieved for large values of λ_1 (remember that λ_1 needs to be positive). We saw in Appendix D that for maximally compact sequences (i.e. sequences that result in the maximum of the dual problem) we have $\lambda_2 = 1/4\; a_0(2\lambda_1)$ (note that $a_0(-q)=a_0(q)$). McLachlan in [27] shows that for large enough values of q, we have

$$\begin{aligned} a_0(q) &= -2q+2q^{\frac{1}{2}}+\frac{1}{4}-\frac{1}{32}q^{\frac{-1}{2}}-\frac{48}{2^7}q^{-1}-\frac{848}{2^{17}}q^{\frac{-3}{2}}-\frac{4,752}{2^{20}}q^{-2}-\frac{126,752}{2^{20}}q^{\frac{-5}{2}}-\cdots \\ &= -2q+2q^{\frac{1}{2}}+\frac{1}{4}-\frac{1}{32}q^{\frac{-1}{2}}+O(q^{-1})\,. \end{aligned} \tag{F.1}$$

Thus, we have for large q,

$$a_0(q) \le -2q+2q^{\frac{1}{2}}+\frac{1}{4}\,. \tag{F.2}$$

After replacing q by $2\lambda_1$ and $a_0(2\lambda)$ by $4\lambda_2$, we have

$$\lambda_2 = \frac{1}{4}a_0(2\lambda_1) \le -\lambda_1+\frac{1}{\sqrt{2}}\sqrt{\lambda_1}-\frac{1}{16}\,. \tag{F.3}$$

Because this set contains the original feasible set of the dual problem, it will give an upper bound on the optimal value of the dual. In other words, $\Delta_n^2 \le \sigma_r^2$, where

$$\begin{aligned} \sigma_r^2 \;=\; &\underset{\lambda_1,\lambda_2}{\text{maximize}} && \alpha\,\lambda_1+\lambda_2 \\ &\text{subject to} && \lambda_2 \le -\lambda_1+\frac{1}{\sqrt{2}}\sqrt{\lambda_1}-\frac{1}{16}\,. \end{aligned} \tag{F.4}$$

It is easy to see that the maximum of (F.4) is achieved for $\lambda_1=\frac{1}{8(1-\alpha)^2}$. Replacing λ_1 in (F.4) and using the fact that $\Delta_n^2 \le \sigma_r^2$, leads to

$$\Delta_n^2 \le \frac{1}{8}\left(\frac{\sqrt{1+\sigma^2}}{\sqrt{1+\sigma^2}-1}-\frac{1}{2}\right)\,. \tag{F.5}$$

$\square$

References

[1] W. Heisenberg, The actual content of quantum theoretical kinematics and mechanics, Physikalishche Z. 43 (1927) 172, (Translated).

[2] M. Vetterli, J. Kovacevic, V. K. Goyal, Foundations of Signal Processing, 2012, http://www.fourierandwavelets.org/.

[3] E. Breitenberger, Uncertainty measures and uncertainty relations for angle observables, Foundations of Physics 15 (3) (1985) 353–364.

[4] K. V. Mardia, P. E. Jupp, Directional Statistics, Wiley, 2009.

[5] J. Prestin, E. Quak, Optimal functions for a periodic uncertainty principle and multiresolution analysis, Proceedings of the Edinburgh Mathematical Society 42 (2) (1999) 225–242.

[6] J. Prestin, E. Quak, H. Rauhut, K. Selig, On the connection of uncertainty principles for functions on the circle and on the real line, Journal of Fourier Analysis and Applications 9 (4) (2003) 387–409.

[7] S. Massar, P. Spindel, Uncertainty relation for the discrete Fourier transform, Phys. Rev. Lett. 100 (2008) 190401.

[8] T. Przebinda, V. DeBrunner, M. Ozaydin, Using a new uncertainty measure to determine optimal bases for signal representations, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1999.

[9] D. L. Donoho, P. B. Stark, Uncertainty principles and signal recovery, SIAM Journal on Applied Mathematics 49 (3) (1989) 906–931.

[10] A. Agaskar, Y. Lu, Uncertainty principles for signals defined on graphs: Bounds and characterizations, in: Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, 2012, pp. 3493 –3496.

[11] Z. Khalid, S. Durrani, P. Sadeghi, R. Kennedy, Concentration uncertainty principles for signals on the unit sphere, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012, pp. 3717 –3720.

[12] R. von Mises, Über die "Ganzzahligkeit" der Atomgewichte und verwandte Fragen, Physikalishche Z. 19 (1918) 490–500.

[13] D. Slepian, Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V:The discrete case 57 (5) (1978) 1371–1430.

[14] A. Brown, P. Halmos, C. Pearcy, Commutators of operators on Hilbert space, Canad. J. Math 17 (1965) 695–708.

[15] E. Schrödinger, About Heisenberg uncertainty relation, Proceedings of The Prussian Academy of Sciences XIX (1930) 296–303.

[16] M. Grant, S. Boyd, cvx: MATLAB software for disciplined convex programming, version 1.21 (2010).

[17] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge University Press, 2004.

[18] A. Shapiro, Rank-reducibility of a symmetric matrix and sampling theory of minimum trace factor analysis, Psychometrika 47 (2) (1982) 187–199.

[19] A. Barvinok, Problems of distance geometry and convex properties of quadratic maps, Discrete & Computational Geometry 13 (1) (1995) 189–202.

[20] G. Pataki, On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues, Mathematics of Operations Research (1998) 339–358.

[21] Z. Luo, W. Ma, A. M. C. So, Y. Ye, S. Zhang, Semidefinite relaxation of quadratic optimization problems, IEEE Signal Processing Magazine 27 (3) (2010) 20–34.

[22] M. Todd, Semidefinite optimization, Acta Numerica 10 (2001) 515–560.

[23] M. Trnovská, Strong duality conditions in semidefinite programming, Journal of electrical engineering 56 (12).

[24] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series 55. Tenth Printing., ERIC, 1972.

[25] A. Erdélyi, W. Magnus, F. Oberhettinger, F. Tricomi, H. Bateman, Higher transcendental functions, Vol. 3, McGraw-Hill New York, 1955.

[26] S. Song Goh, C. A. Micchelli, Uncertainty principles in hilbert spaces, Journal of Fourier Analysis and Applications 8 (2002) 335–374.

[27] N. W. MacLachlan, Theory and Applications of Mathieu Functions, Dover Publications, 1964, p. 232.